





```

Db 150 GGTAATCTAGTACCATGCTGTTACTCTCAGGCATGGATGAATAAATGTTATAGCTG 91
Qy 3060 ca 3061
Db 90 CA 89

RESULT 3
LOCUS AW874082 593 bp mRNA EST 22-MAY-2000
DEFINITION hq82g12.x1 NCI_CGAP_Thy3 Homo sapiens cDNA clone IMAGE:3125926 3'
similar to gb:J02970 THYROID PEROXIDASE 2 PRECURSOR (HUMAN);, mRNA
sequence.
ACCESSION AW874082
VERSION AW874082.1 GI:8008135
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 593)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Klaus Kaserer, M.D., Chris Moskaluk, M.D.,
Ph.D., Michael R. Emmert-Buck, M.D., P.h.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 371.
FEATURES
source
1..593
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3125926"
/clone_lib="NCI_CGAP_Thy3"
/tissue_type="follicular carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: thyroid; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library constructed by Life
Technologies."
BASE COUNT 134 a 157 c 153 g 148 t 1 others
ORIGIN
Query Match 15.2%; Score 465.8; DB 10; Length 593;
Best Local Similarity 97.3%; Pred. No. 9.3e-90;
Matches 473; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2576 taggagacatggagaaacctgcgtagactccggaggctccctcggtgacttgatct 2635
Db 570 TAAGAACCAGTGGAGAACCTGGTACACTCTCGGAGGTTCCCTCGTGAGACTTGGATCT 511
Qy 2636 caatgcctgcgtcgtctcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgt 2695
Db 510 CCATGTCGCTGGTTCCTCTGCTGATCGGAGGCTTCGACGCTCTCACCTCGACGCTGATNT 451
Qy 2696 gcaggtggagacacactgcactgaataatcacactgccatctcggagacagcgaggaa 2755
Db 450 GCAGGTGGACACACACTGGACATTAATTCACACTGCCATCTCGGACAGCGGAGGAA 391
Qy 2756 ctcccagctgagatcggaaagacacacagcgcgtagggacctcacgcagcggccgcag 2815
Db 390 CTCCCGAGCTGAGATCGGAAAGACACAGCGCCGTAGGGACCTCACCGCAGCGGCGCAG 331

```

```

Qy 2816 ctcaggactcggagcagagagtgctggtgagtggaagccgggatactcacagcgtgcgga 2875
Db 330 CTCAGGACTCGGAGCAGGAGAGTGTGGATGGAAGCCGGGATACCTCACAGGCTGCCGA 271
Qy 2876 gagccctctgagggcaaaagtggcaggacactgcagacagcttcattgttcccaaatcac 2935
Db 270 GAGCCCTCTGAGGCAAAAGTGGCAGGACACTGCAGAACAGGCTTCATGTGCCAAATCAC 211
Qy 2936 cgtacgactctttccaaacacagcagcaaatcgaaatcagcagcagcagcgtgtttcccaa 2995
Db 210 GGTACGACTCTTTTCCAAACACAGCAATCCGAATCAGCAGGACGACTGTTTCCCAA 151
Qy 2996 cacgggtaaatctagtcacatgctgtagttactctcagcagcagcagcagcagcagcagc 3055
Db 150 CACGGGTAAATCTAGTACCATGTCGTAGTACTCTCTCAGCATGGATGAATAAATCTTATA 91
Qy 3056 gctgca 3061
Db 90 GCTGCA 85

RESULT 4
LOCUS AW874069/c
DEFINITION hq82ell.x1 NCI_CGAP_Thy3 Homo sapiens cDNA clone IMAGE:3125900 3'
similar to gb:J02970 THYROID PEROXIDASE 2 PRECURSOR (HUMAN);, mRNA
sequence.
ACCESSION AW874069 595 bp mRNA EST 22-MAY-2000
VERSION AW874069.1 GI:8008122
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 595)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Klaus Kaserer, M.D., Chris Moskaluk, M.D.,
Ph.D., Michael R. Emmert-Buck, M.D., P.h.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Possible reversed clone: polyT not found
Seq primer: -400P from Gibco
High quality sequence stop: 412.
FEATURES
source
1..595
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3125900"
/clone_lib="NCI_CGAP_Thy3"
/tissue_type="follicular carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: thyroid; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library constructed by Life
Technologies."
BASE COUNT 124 a 175 c 163 g 132 t 1 others
ORIGIN
Query Match 14.1%; Score 434; DB 10; Length 595;
Best Local Similarity 80.9%; Pred. No. 6.2e-83;
Matches 577; Conservative 0; Mismatches 6; Indels 130; Gaps 1;

```



EST.  
SOURCE  
ORGANISM

Db 143 AGCCCCGAGCTGGAAACCCCGGCTTCTTTGATCAACGGGTTCACACTGCCCGGTCGGG 84

Qy 704 aggtgacaaagacatgtcattcaagtattcaaatgaggtgtcacagatgatgcacggctatt 763

Db 83 AGGTGACAAGACATGTCATTCAAGTTTCAAATGAGGTGTGCACAGAT--TGACCCTATT 26

Qy 764 ctgacctccctgatggcatgggggaca 788

Db 25 CTGACCTCCTGACGCTCGGGGTTCCA 1

RESULT 7  
BI004574/c  
LOCUS  
DEFINITION  
MR4-HN0054-070301-002-g06 HN0054 Homo sapiens cDNA, mRNA sequence.  
ACCESSION  
BI004574  
VERSION  
BI004574.1 GI:14408648  
KEYWORDS  
EST.  
SOURCE  
HUMAN.

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 (bases 1 to 457)

AUTHORS  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

TITLE  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE  
20202663

COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV1&t2-QV1-HB0032-  
111200-560-c07&t3=2000-12-11&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 6  
High quality sequence stop: 443.  
Location/Qualifiers  
1..443  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="HB0032"  
/dev\_stage="Adult"

FEATURES  
source  
1..457  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="HN0054"  
/dev\_stage="Adult"  
/note="Organ: head\_normal; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
86 a 101 c 132 g 124 t

BASE COUNT  
ORIGIN

Query Match 13.7%; Score 419.8; DB 11; Length 443;  
Best Local Similarity 98.0%; Pred. No. 6.6e-80;  
Matches 436; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Qy 344 ttccccgacgacagataatggaacatcataacacagcgatgaaaagaagtcacc 403

Db 443 TTCCCCGACGACGAGATAAATGGAACATCATCAACGCGATGAAAAGAATGCAACC 384

Qy 404 tgaasactcaaacatcacagctccaacggatgctttatcagaagatctgctgagcatca 463

Db 383 TGAARAACTCAACAATCACAGATCAACAGGATGCTTTATCAGAAGATCTGTGACATCA 324

Qy 464 ttgcaaacatgtctgtagtctcccttacatgctgcccccacaaatgcccaacactgcc 523

Db 323 TTGCAAAACATGTCTGGATGCTCCCTTACATGCTGCCCCCACAAATGCCAACACTTGCC 264

Qy 524 tgycgaaaaatacaggcccatcacaggagcttgcaaacacagagacacccccagatggg 583

Db 263 TGCGCAACAAATACAGCCCATCAGAGGAGCTTGAACAACAGAGACACCCCGATGGG 204

Qy 584 gcgcctcccaacagccctggcacatgctccctccagctctatgagacgcttcagtc 643

Db 203 GCOCCTCCAAACACGGCCCTGGCACGATGCTCCCTCCAGTCTATGAGACGCTTCAGTC 144

Qy 644 agccccgaggtggaacccccggctctttatacaacgggttcccactgcccccggtccggg 703

QY 2147 agcaactccctgtctcgggtcattgtgaaacacactggtcctcaccagggtgcccattgatg 2206  
 Db 397 AGCACTCCCTGTCTCGGTCATCTGTGCACACACTGGCTCACCAGGGTGCCCATGGATG 338  
 QY 2207 ccttcaagtccggaataatcccggaagactttgagttgtagacacatcacatcagtgatga 2266  
 Db 337 CCTTCCAAAGTCGCAATATCCCTGAAGACTTTGAGTCTGTGTGACAGATCCCTGGCATGA 278  
 QY 2267 acctgagggcctggagggaacaccttctcctcaagacacagactggtggttcccccagagagc 2326  
 Db 277 ACCTGGAGGCTTGGAGGGAACCTTTCTCAACACGACAGTGTGGCTTCCAGAGAGCG 218  
 QY 2327 tgagaaatggggaacttctgacactgtgagagtcctctggaggcgctgctgtgtattcct 2386  
 Db 217 TGGAGAAATGGGACTTTGTGCACTGTGAGAGTCTGGAGGCGCTGCTGGTGTATTCTT 158  
 QY 2387 gccggcacaggatagctcccaagggcgagcagcagctcacttaccacccaggaagatggg 2446  
 Db 157 GCGGGCAGGGTATGAGCTCCAAAGGCGGAGCAGCTCACTTCCACCCCGAAGGATGGG 98  
 QY 2447 attccagctccctctcgtcaaaagtgtgaacagtgtagcagcagtggtgcccacccccct 2506  
 Db 97 ATTTCCAGCTCCCTCTGCCAAGATGTGAACGAGTGTGCAGACCTGCTGCCACCCCGC 39  
 QY 2507 gcaacgctctcgtgaggtgcaagaacacacacacacacacacacacacacacacacac 2545  
 Db 40 -CTGCCACCTTCGAGGTGCAGAAACACCATGATCAGT 3

RESULT 8  
 AW873387  
 LOCUS  
 DEFINITION h19c12.x1 NCI-CGAP\_Thy8 Homo sapiens cDNA clone IMAGE:3011158 3', similar to gb:J02970 THYROID PEROXIDASE 2 PRECURSOR (HUMAN);, mRNA sequence.  
 ACCESSION AW873387  
 VERSION AW873387.1 GI:8007440  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 422)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-re@mail.nih.gov  
 Tissue Procurement: Klaus Kaserer M.D., Bruno Niederle M.D., Mike Emmert-Buck M.D. Ph.D., Vlado Knezevic M.D.  
 CDNA Library Preparation: Krizman Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL, send email to: info@image.lnl.gov  
 Seq primer: -40UP from Gibco.  
 Location/Qualifiers  
 1. 422  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3011158"  
 /clone\_lib="NCI-CGAP\_Thy8"  
 /tissue\_type="normal epithelium"  
 /lab\_host="DH10B"  
 /note="Organ: thyroid; Vector: pAMP10; mRNA made from normal thyroid epithelium, cDNA made by oligo-dT priming. Non-directionally cloned into UDG sites. Size-selected on agarose gel, average insert size 500 bp. Primary library. CDNA Library Preparation: David B. Krizman, Ph.D. REFERENCE: Krizman et al. (1996) Cancer Research

BASE COUNT 92 a 122 c 140 g 68 t  
 ORIGIN

Query Match 12.7%; Score 389.2; DB 10; Length 422;  
 Best Local Similarity 97.1%; Pred. No. 2.4e-73;  
 Matches 408; Conservative 0; Mismatches 8; Indels 4; Gaps 1;

QY 2503 cctctgccacgcctctgagagtgagaaacacacacacacacacacacacacacacacacac 2562  
 Db 7 CGCGCGCGCTCGAGCGCGAGGTGCAGAAACACCAAGAGCGGCTTCAGTGTCTCTGCGCG 66  
 QY 2563 gacctctacaggttagagacgatgggagaaacctgctagactccggagggtccctcctcgg 2622  
 Db 67 GACCCCTACGAGTTAGGAGACGATGGGAGAACCTCGCTAGACTCCGGGAGGCTCCTCTCG 126  
 QY 2623 gtgacttgatctccatctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2682  
 Db 127 GCGACTTGGATCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 186  
 QY 2683 tcgacgtgatttcaggtgacacgacactgacactaaatcacacactgccatctcgag 2742  
 Db 187 TCGACGTTGATTTGAGGTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246  
 QY 2743 acagcgaggagaaacctccgagctgagatcggaagcaccagcgctgaggagacacacac 2802  
 Db 247 ACAGCGGAGGAACCTCCGAGCTGAGATGCGAAGACACAGGCGCTAGGGACCTCACCG 306  
 QY 2803 cagcgggcccagctcagactcagcagcagcagcagcagcagcagcagcagcagcagcag 2862  
 Db 307 CAGCGGCGCGCAGCTCAGGACTCGGA-----GAGAGTGTGGATGGAGGCGGGGATACT 362  
 QY 2863 cacaggtcgccagacgacctctgaggaagcagcagcagcagcagcagcagcagcagcagc 2922  
 Db 363 CACAGGCTGCGAGAGCGCTCTGAGGCGCAAGTGGCAGGACACTGCAGAACAGCTTCATG 422

RESULT 9  
 AA384625  
 LOCUS  
 DEFINITION EST98168 Thyroid Homo sapiens cDNA 5' end similar to thyroid peroxidase, mRNA sequence.  
 ACCESSION AA384625  
 VERSION AA384625.1 GI:2036944  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 340)  
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagan,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Morenno-Falancan,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,K.K., Kozak,D.L., Kunsch,C., Hunglun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.  
 TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence  
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)  
 MEDLINE 96026280  
 COMMENT Other\_ESTs: THC123110

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tldb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1. .340

/organism="Homo sapiens"

/db\_xref="ATCC (inhost):188788"

/db\_xref="taxon:9606"

/clone\_lib="Thyroid"

/dev\_stage="adult"

/note="Organ: thyroid gland; Vector: pBluescript SK-;

Site\_1: EcoRI; Site\_2: XhoI"

91 a 97 c 97 g 55 t

BASE COUNT  
ORIGIN

Query Match 11.0%; Score 337.4; DB 10; Length 340;  
Best Local Similarity 99.7%; Pred. No. 3.1e-62;  
Matches 338; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2701 tgacacgactgacactaaatccacactgccatctcgagacagcgaggaactccc 2760  
|||||  
Db 1 TGGACAGCACTGGCACTAAATCCACTGCCATCTCGGAGACAGCGGAGGAACCTCC 60  
|||||  
QY 2761 gagctgagatgcggaagcaccagccgctgaggaacctcaccgagcgccgagctcag 2820  
|||||  
Db 61 GAGCTGAGATGCGGAAGCACACAGCGCGTAGGACCTCACCGAGCGCGCGAGCTCAG 120  
|||||  
QY 2821 gactcgagcagagagtgctggtggaagccgggatactcacagctgcgcgagagcc 2880  
|||||  
Db 121 GACTCGGAGCAGAGAGTGTGGGATGGGAAGCGCGGATACTCACAGGCTGCCGAGAGCC 180  
|||||  
QY 2881 ctctgagggcaagtgcgagacacagcagcttcattgtttcccaaatcacctac 2940  
|||||  
Db 181 CTCTGAGGGCAAGTGGCAGGACACTGCAGAACAGCTTCATGTTCCCAAAATCACCGTAC 240  
|||||  
QY 2941 gactctttcccaaacacaggaatcggaatcagcagagcagctgtttcccaaacagg 3000  
|||||  
Db 241 GACTCTTTCCAAACACAGGCAATCCGAATCAGAGGAGCAGTGTTCCTCCCAACAGG 300  
|||||  
QY 3001 gtaaatctagttaccatgtgttagttactctcaggcatgg 3039  
|||||  
Db 301 GTAAATCTAGTACCATGTCGTAGTTACTCTCAGGCATGG 339  
|||||

RESULT 10

BG187512 383 bp mRNA EST 21-APR-2001  
LOCUS RST6635 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
DEFINITION BG187512  
ACCESSION BG187512  
VERSION BG187512.1 GI:13709327  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 383)  
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Dahl,T., Thornton,M., Ramachandran,R., Whittington,J.,  
Lerner,L., Krashoc,D., McElligott,K., Clark,S., Mays,R., Smith,E.,  
Veloso,N., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J.  
and Ducar,M.  
TITLE Creation of Genome-wide Protein Expression Libraries using Random  
Activation of Gene Expression  
JOURNAL Nat. Biotechnol. 19 (5), 440 (2001) In press  
COMMENT

COMMENT Contact: Scott J. Cain

Athersys, Inc.

3201 Carnegie Ave, Cleveland, OH 44115, USA

Tel: 216 431 9900

Fax: 216 361 9596

Email: scain@atersys.com

High quality sequence stop: 383.

Location/Qualifiers

1. .383

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="Athersys RAGE Library"

/cell\_line="HT1080"

/note="See 'Creation of Genome-wide Protein Expression

Libraries using Random Activation of Gene Expression',

Nature Biotechnology, in press. Note that even though the

cell type indicated is HT1080, since a random activation

method was used, these sequence tags are not necessarily

expressed in HT1080 under normal circumstances."

114 a 98 c 102 g 65 t 4 others

BASE COUNT  
ORIGIN

Query Match 10.8%; Score 333; DB 11; Length 383;  
Best Local Similarity 97.7%; Pred. No. 2.8e-61;  
Matches 347; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 2719 aaattccacactgccatctcgagacagcgaggaactcccgagctgag-atgcgga 2777  
|||||  
Db 29 AATNCCACACTGCCATCTCGGNAACGGCGGAGGAACCTCCGAGCTGAGATGCGGAAA 88  
|||||  
QY 2778 gcaccagggccgtagggacctcaccgagcgagcgccgagctcaggactcgagcagagag 2837  
|||||  
Db 89 GCACCAGGCCCTAGGAGACTCACCGCAGCGGCCGAGCTCAGGAGCTCGGAGCAGGAGAG 148  
|||||  
QY 2838 tgctgggagtgaagccgggatactcacaggtcgcgagagccctctgaggcgaagtgg 2897  
|||||  
Db 149 TCCTGGGATGGAAGCGCGGATACTCACAGGCTGCGGAGAGCCCTCTGAGGCAAGTGG 208  
|||||  
QY 2898 cagacactgcagaacagcttcattgttcccaaatcacgctacgactctttcccaaac 2957  
|||||  
Db 209 CAGGACACTGCAGAACAGACTTCATGTTCCCAAAATCACCGTACGACTCTTTCCCAACAC 268  
|||||  
QY 2958 aggcacaaatcggaatcagcagcagcagctgtttcccaaacacggttaaatctagtaccatg 3017  
|||||  
Db 269 AGGCAAAATCGGAATCAGCAGGACGACACTGTTTCCCAACACGCGTAAATCTAGTACCATG 328  
|||||  
QY 3018 tcgtagtactctcaggcatggatgaataaaatgttatagctgcacacacacacacacac 3072  
|||||  
Db 329 TCGTAGTTACTCTCAGGCATGGATGAATAAATGTTATAGCTGAAAAAATGAAAAA 383  
|||||

RESULT 11

BG213066 365 bp mRNA EST 21-APR-2001  
LOCUS RST32668 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
DEFINITION BG213066  
ACCESSION BG213066  
VERSION BG213066.1 GI:13734753  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 365)  
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Dahl,T., Thornton,M., Ramachandran,R., Whittington,J.,  
Lerner,L., Krashoc,D., McElligott,K., Clark,S., Mays,R., Smith,E.,  
Veloso,N., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J.  
and Ducar,M.  
TITLE Creation of Genome-wide Protein Expression Libraries using Random  
Activation of Gene Expression  
JOURNAL Nat. Biotechnol. 19 (5), 440 (2001) In press  
COMMENT Contact: Scott J. Cain

Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: [scaine@atersys.com](mailto:scaine@atersys.com)  
High quality sequence stop: 365.

<b>FEATURES</b>	<b>SOURCE</b>
-----------------	---------------

COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&t2=MR3-RT0999-070201-005-all&t3=2001-02-07&t4=1>)

FEATURES  
SOURCE

```

FEATURES
source
    1..328
        Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="HT0999"
            /dev_stage="Adult"
            /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                Site_2: SmaI; A mini-library was made by cloning products
                derived from ORESTES PCR (U.S. Letters Patent application
                No. 196,716 - Ludwig Institute for Cancer Research)
                profiles into the pUC 18 vector. Reverse transcription of
                tissue mRNA and cDNA amplification were performed under
                low stringency conditions."
            67 a          90 c      97 g      74 t
BASE COUNT
        ATGCTT

```

RESULT	12
BG993369	
LOCUS	BG993369 328 bp mRNA EST
DEFINITION	MR3-HT0999-070201-005-all HT0999 Homo sapiens cDNA, mRNA sequence.
	13-JUN-2001

**REFERENCE**  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 328)  
Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663

**TITLE**

	Query Match	10.1%	Score 310;	DB 11;	Length 328;
	Best Local Similarity	100.0%;	Pred. No. 2.4e-56;		
	Matches 310;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1947	gatcctggactgtacaagcatcctgacaacatcgatgtctgctggtaggacttaactga	2006		
Db	19	GATCCCGGACTTGTACAAAGCATCCTGACAACTCGATGCTGCTGGAGGCTTAGCTGA	78		
QY	2007	aaacttcctcccaagggctcggacaggggccctgtttgcctgtctcattgggaagcagat	2066		
Db	79	AAACTCCCTCCCCAGGGCTCGGACAGGGCCCCGTGTTGCCCTGTCATTGGGAAGCAGAT	138		
QY	2067	gaaggctctcgggacggtgactggtttggtgggagaaacagccacgtcttcacgagtgc	2126		
Db	139	GAAGGCTCTCGGGAGCGGTGACTTGGTTTTTGGTGGGAGACAGCCACGCTCTTCACGGATGC	198		
QY	2127	acagagcctgagctggagaagcactccctgtctcgggtcactctgtgacaacactggcct	2186		
Db	199	ACAGAGCGGTGAGCTGGGAAGCACATCCCTGTCTCGGGTCACTGTGACAACTGGCCT	258		
QY	2187	caccaggggtgccatggatgccttccaaagtccggcaaatccccqaaagacttttagtcttg	2246		
Db	259	CACCAGGTTGCCATGGATGCCCTTCCAAGTCGGCAATTCCTCCCGAAGACTTTCAGCTTGG	318		
QY	2247	tgacagcatc	2256		
Db	319	TGACAGCATC	328		

RESULT	13
AK018070	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

**TITLE** Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

**MEDLINE** 20203663



[illegible]

```

QY 2155 ctgtctcgggtcatctgtgacacacacgctccaccaggtgcccagtggtgagctctccaa 2214
Db 848 TTCTCCCGCTGCTGTGTGACAACTGCGATCGCAAGGTCCCACTGACACCGGTTCCAG 907
QY 2215 qtcgcaaatcccggaacatttgagttgttgacagacatcatcggcatgaacctggag 2274
Db 908 GCTATGCGCTACCCCGACGCTTGTGGACTGCTCTCAATGATAAGCTGGACCTCTCA 967
QY 2275 gctggg 2280
Db 968 CCTGG 973

RESULT 14
AW874128/c
LOCUS AW874128 339 bp mRNA EST 22-MAY-2000
DEFINITION hq85a04.x1 NCI_CGAP_Thy3 Homo sapiens cDNA clone IMAGE:3126126 3'
similar to gb:J02970 THYROID PEROXIDASE 2 PRECURSOR (HUMAN);, mRNA
sequence.
ACCESSION AW874128
VERSION AW874128.1 GI:8008181
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 339)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Ph.D., Michael R. Emmert-Buck, M.D., P.h.D.
Tissue Procurement: Klaus Kaserer, M.D., Chris Moskaluk, M.D.,
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 218.
Location/Qualifiers
1. 339
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3126126"
/clone_lib="NCI_CGAP_Thy3"
/tissue_type="follicular carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: thyroid; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.4 kb. Library constructed by Life
Technologies."
BASE COUNT 57 a 89 c 87 g 95 t 11 others
ORIGIN
source

Query Match 8.6%; Score 263.2; DB 10; Length 339;
Best Local Similarity 87.9%; Pred. No. 2.6e-46;
Matches 291; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 2736 ctggagacagcgagggaactcccgagctgagatgcggaaagcacacagccctaggagc 2795
Db 331 CTTTCGAACAGGCGGAGAACTCCCNNGNATAGATCGGAAGCANCAGCGGAGGAC 272
QY 2796 ctacacgagcgaggcgagctcagactcggag-caggagagtgctggatggaagcc 2854
Db 271 CTCNNCCGAGCGGCGCAGCTCAGGACTCGGNACAGGAGAGTGTGGATGGAAGGCC 212
QY 2855 gggatactcacaggtgccgagagccctctgagggcaaaagtgcgagacactgcagaaca 2914

```

```

Db 211 GGGATACCTCAGAGCTGCCGAGAGCCCTCTTGAGGCAAAAGTGGCAGACACTNCAGAAACA 152
QY 2915 gcttcatttcccaaaatcacgcgtacgactttttccaaacacagcgaaatcggaataca 2974
Db 151 GCTTCATGTTCCCAAAATCATCCGTCAGCACTCTTTTCCAAACACACAGGCAAAATCCGAAATCA 92
QY 2975 gcaggacgactgttttcccaacacggtgtaatacttagtaccatgtcgttagttactctcagg 3034
Db 91 GCAGGACGACTGTTTTCGCAACACGGGTAAATCTAGTACCATGTCTGCTAGTACTCTCAGG 32
QY 3035 catgagatgaataaatgttatagctgcacaaaaa 3065
Db 31 CATGGATGAATAAATGTATAGCTGCAAAAAA 1

RESULT 15
AA384833
LOCUS AA384833 261 bp mRNA EST 21-APR-1997
DEFINITION EST98587 Thyroid Homo sapiens cDNA 5' end similar to thyroid
peroxidase, mRNA sequence.
ACCESSION AA384833
VERSION AA384833.1 GI:2037317
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 261)
AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult
, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White
, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A.,
Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald
, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A.,
Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkley, P.S., Jr., Kelley, J.M.,
Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Saudke, D.M., Shirley, R.,
Snall, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferlie, A., Fischer, C., Hastings, G.A., He, W.W.,
Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,
Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L.,
Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon
, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and
Venter, J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT Other ESTs: THC123110
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 261
/organism="Homo sapiens"
/db_xref="ATCC (inhost):189160"
/db_xref="taxon:9606"
/clone_lib="thyroid"
/dev_stage="adult"
/notes="Organ: thyroid gland; Vector: pBluescript SK-;
Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 69 a 73 c 72 g 45 t 2 others
ORIGIN

```

Query Match 8.4%; Score 257.4; DB 10; Length 261;  
 Best Local Similarity 98.9%; Pred. No. 4.5e-45;  
 Matches 258; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2779 caccagggcgttaggacccaccagcgccgagctcaggaactcggagcagagagt 2838  
 Db 1 CACCAGGCCGTAGGNCCTCACCGACGGCGCGCAGCTCAGGACTCGGAGCAGGAGGT 60

QY 2839 gctgggatggaaggccgggatactcacagggctgcccagagagccctctgagggcaagtggc 2898  
 Db 61 GCTGGGATGGAGGCCGGGATACTCACAGGCTGCCGAGAGCCCTCTNAGGGCAAGTGGC 120

QY 2899 aggaactgcagaacagcttctatgtttcccaaaatcccgtaagactcttttccaaacaca 2958  
 Db 121 AGGACACTGCAGAACAGCTTCATGTTCCCAAAATCACCGTACGANTCTTTTCCAAACACA 180

QY 2959 gcaaatcgaaatcagcagagcagactgttttcccaaacacgggtaaatctagtaccatgt 3018  
 Db 181 GGCAATCCGAATCAGCAGGACGACTGTTTCCCAACACGGGTAAATCTAGTACCATGT 240

QY 3019 cgtagtactctcaggcatgg 3039  
 Db 241 CGTAGTTACTCTCAGGCATGG 261

Search completed: November 26, 2001, 12:26:57  
 Job time: 3611 sec